DOCUMENT RESUME

ED 459 430 CS 014 560

AUTHOR	Paris, Scott G.; Paris, Alison H.; Carpenter, Robert D.
TITLE	Effective Practices for Assessing Young Readers. CIERA Report.
INSTITUTION	Center for the Improvement of Early Reading Achievement, Ann Arbor, MI.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	CIERA-3-013
PUB DATE	2001-07-17
NOTE	23p.
CONTRACT	R305R70004
AVAILABLE FROM	CIERA/University of Michigan, 610 E. University Ave., 1600 SEB, Ann Arbor, MI 48109-1259. Tel: 734-647-6940; Fax: 734-763-1229. Web site: http://www.ciera.org.
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Beginning Reading; *Informal Reading Inventories; Primary Education; *Reading Comprehension; *Reading Skills; *Student Evaluation; Teacher Surveys
IDENTIFIERS	Reading Fluency

ABSTRACT

This report addresses how reading comprehension can be assessed when children have limited decoding skills. Successful teachers use reading assessments for many purposes. A Center for the Improvement of Early Reading Achievement (CIERA) survey of 496 teachers in "beat the odds" schools revealed that they use many informal reading tasks to assess children's skills, knowledge, and fluency on a daily basis. They use "internal" assessments under their control more often than "external" assessments created by publishers or mandated by others. Paradoxically, teachers believe that administrators give more attention to external, high-stakes assessments than the internal assessments of daily work that are linked to instruction. It is also paradoxical that teachers report their training for internal assessments as "Good" but their training to use external assessments was only "Fair." This report outlines the benefits of a developmental approach to early reading assessment based on teachers' internal assessments that are tailored selectively for children and the curriculum. (Contains 4 tables, guidelines for assessment with children 3-8 years of age, and 26 references.) (NKA)

Reproductions supplied by EDRS are the best that can be made from the original document.

ED 459 430

CIERA REPORT

Effective Practices for Assessing Young Readers

Scott G. Paris
Alison H. Paris
Robert D. Carpenter
University of Michigan

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

CIERA REPORT #3-013
July 17, 2001

CIERA
Center for the Improvement of Early Reading Achievement

CS 014 560

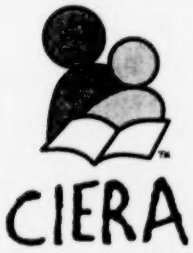
Effective Practices for Assessing Young Readers

CIERA REPORT #3-013

Scott G. Paris, Alison H. Paris, and Robert D. Carpenter
University of Michigan

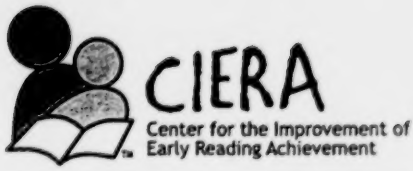
CIERA Inquiry 3: Policy and Profession
How can we assess comprehension when children have limited decoding skills?

Successful teachers use reading assessments for many purposes. Our CIERA survey of teachers in "beat the odds" schools revealed that they use many informal reading tasks to assess children's skills, knowledge, and fluency on a daily basis. They use "internal" assessments under their control more often than "external" assessments created by publishers or mandated by others. Paradoxically, teachers believe that administrators give more attention to external, high-stakes assessments than the internal assessments of daily work that are linked to instruction. It is also paradoxical that teachers report their training for internal assessments as "Good" but their training to use external assessments was only "Fair." We outline the benefits of a developmental approach to early reading assessment based on teachers' internal assessments that are tailored selectively for children and the curriculum.



Center for the Improvement of Early Reading Achievement
University of Michigan – Ann Arbor

July 17, 2001



University of Michigan School of Education

610 E University Ave., Rm 1600 SEB
Ann Arbor, MI 48109-1259

734.647.6940 voice
734.615.4858 fax
ciera@umich.edu

www.ciera.org

©2001 Center for the Improvement of Early Reading Achievement.

The report described herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305R70004, as administered by the Office of Educational Research and Improvement, U.S. Department of Education. However, the contents of the described report do not necessarily represent the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment or the National Institute on Early Childhood Development, or the U.S. Department of Education, and you should not assume endorsement by the federal government.

Effective Practices for Assessing Young Readers

Scott G. Paris, Alison H. Paris, Robert D. Carpenter
University of Michigan

Assessment is a vital part of successful teaching because instruction needs to be calibrated according to students' knowledge, skills, and interests. Tests, quizzes, and performance evaluations help teachers identify developmentally appropriate instruction. Effective instruction challenges children because it is on the edge of their independent abilities, the "zone of proximal development" in Vygotsky's terms. Effective instruction may also be fun, inspirational, and motivating. Most importantly, effective instruction is shaped by assessment because teachers use their knowledge about students to select materials based on interest and difficulty, and to group children based on collaborative work habits. Some of these decisions may not be regarded as "assessment" in a traditional sense, but they illustrate how teachers use their informal knowledge about children to guide their classroom instruction.

Successful teachers use reading assessments for many purposes. They may use informal assessments at the start of the school year to become familiar with students' fluent reading. They may use skill tests to diagnose strengths and weaknesses. They might observe decoding and comprehension strategies during daily reading. They might design self-assessments so students can monitor their own progress. They might use journals to monitor changes in children's handwriting, reading interests, and phonetic approximations to words. Of course, they prepare their students for high-stakes tests, too. With many different kinds of assessments used for many different purposes, today's teachers need to be knowledgeable about when and why to use the various tools available to them (Shepard, 2000).

Internal and External Assessments

Some reading assessments are informal, frequent, and tied to curriculum and daily instructional routines in the classroom. For example, assessments of children's daily oral language, listening, and question-answering during group reading may be made through teachers' observation. Other assessments may be more structured, such as spelling tests, weekly quizzes, journal writing, reports, and projects; but they are all under the control of the teacher and embedded in the curriculum. We refer to these assessments as

"internal" because they are designed, selected, and used by teachers according to the needs of their children. Internal assessments are used to make decisions about instruction and to report progress to parents. In contrast, "external" assessments are designed, selected, and controlled by another person or group—commercial publishers, district administrators, or state policymakers. Typical examples of external assessments include standardized and commercial reading tests. External assessments occur less frequently than internal assessments, but they usually have greater importance, more authority, and higher stakes attached to them. External assessments have been used as indicators of both the educational achievement of students and the quality of instruction in schools. Although external assessments are used most often in grade 4 and beyond, there has been an increasing tendency to use external reading assessments in K–3 classrooms. Thus, we will briefly discuss the impact of high-stakes tests before we examine the variety of internal assessments used by successful teachers.

High-Stakes Testing

Although we believe that the primary function of assessment is to promote teaching and learning in the classroom, assessment has increasingly become a means of enforcing educational accountability, and it reaches beyond the classroom. Commercial tests are used to measure mastery of the curriculum, norm-referenced tests are used to compare students to national expectations, and criterion-referenced tests are used to evaluate the attainment of state-endorsed standards of achievement (National Commission on Testing and Public Policy, 1990). During the past 20 years, there has been a steady increase in the use of standardized tests as accountability measures (Linn, 2000; Madaus & Tan, 1993). There has also been a parallel increase in concerns expressed about the liabilities of increased testing (Shepard, 2000). Some worry that the curriculum has been narrowed (Haertel, 1989); some worry that teachers are being judged inappropriately on the bases of standardized tests (Smith, 1991); and some worry that increased testing has negative effects on students' learning and motivation (Paris, Turner, Lawton, & Roth, 1991; Paris, 2000). Teachers and administrators are worried about being judged inappropriately on the bases of standardized tests (Paris & Urdan, 2000; Smith, 1991). The issue has such profound political and educational implications for reading that the International Reading Association (1999) and the American Educational Research Association (2000) published position papers pointing out the potential problems with high-stakes testing.

Several researchers have examined the impact of high-stakes testing on teachers. For example, Haladyna, Nolen, and Haas (1991) studied Arizona teachers' views of the state-mandated high-stakes test and found that many teachers thought the test was unfair to minority and ESL students. Nolen, Haladyna, and Haas (1992) reported that many teachers engaged in inappropriate or unethical testing procedures because of pressure to produce high test scores with their students. In surveying teachers about the state-required test in Michigan, Urdan and Paris (1994) found that many Michigan teachers were frustrated by external pressures to "teach to the test" and

angry that the tests were used to evaluate teachers' effectiveness. Hoffman, Assaf, Pennington, and Paris (2001) found that teachers in Texas felt coerced to teach skills relevant to the TAAS (Texas Assessment of Academic Skills) to the exclusion of other subjects. Many Texas teachers, like the Michigan and Arizona teachers, believed that the standardized tests were unfair to minority and ESL students. Shepard (1991) pointed out that teachers have little control over the policies proscribing accountability through "high-stakes" tests. A growing number of educators regard such tests of reading as "fragile evidence" of children's reading accomplishments (Murphy, Shannon, Johnston, & Hansen, 1998). The pressure for accountability through testing, coupled with the lack of involvement of teachers in setting policies, has left many teachers frustrated with the growing influence of externally-imposed testing on their professional practices.

Assessment and accountability have become the centerpieces of many educational reforms, with direct implications for teachers' daily practices. In addition to high-stakes tests, teachers are increasingly required to use new assessment tools in their classrooms. For example, teachers in the last decade have been encouraged or required to collect work samples, student portfolios, and informal assessments that are aligned with the curriculum. They often design new district-level tests and report cards. It seems paradoxical that teachers are more involved with "low-stakes" assessments, but they are still judged publicly by the results of high-stakes tests. Because it is evident that teachers are being asked to become more proficient in designing, administering, and interpreting a variety of educational assessment tools in their classrooms, it is important to provide them with the knowledge and training they need to use assessments prudently and effectively in their classrooms.

In the remaining parts of this chapter we describe effective assessment practices that K-3 teachers use in their classrooms. We begin with a report of a large national survey of teachers in outstanding schools to learn about the kinds of reading assessment tools they used and the purposes of these tools. We also report teachers' opinions about the impact of various kinds of assessments on children, parents, and administrators. Next, we go beyond the survey to outline a developmental approach to assessment for young children. We conclude with a discussion of typical assessment problems that teachers must solve and a list of recommendations for effective assessment.

What Teachers Say About Reading Assessment

The survey was designed to collect teachers' perceptions of assessment—specifically, reading assessment in early elementary grades. We wanted to ask successful teachers what kinds of reading assessments they use for what purposes, so that a collection of "best practices" might be available as models for other teachers. We also wanted to know if teachers felt adequately trained to administer these assessments and what they believed to be the impact of various assessments on students, parents, teachers, and administrators. Thus, we decided to survey elementary teachers who taught in "beat the odds" schools to determine their practices and views.

"Beat-the-odds" schools across the nation were defined as schools with a majority of students who qualified for Title I programs and had a mean school test score on some standardized measure of reading achievement that was higher than the average score of other Title I schools in that same state. In most cases, the selected schools also scored above the state average for all schools. Candidate schools were selected from a network of CIERA partner schools, as well as from annual reports of outstanding schools in 1996, 1997, and 1998, as reported by the National Association of Title I Directors. In April 1998, survey packets were sent to more than 400 nominated schools across the nation. Each packet contained a principal survey and seven teacher surveys, in addition to directions to the principal to select seven "key" teachers from grades K-3 to complete the teacher surveys. Approximately 700 K-3 staff from 140 schools responded to the surveys and, specifically, to the questions on assessment practices. The final sample of 504 teachers was established by omitting reading specialists, teachers who taught multiple grades, and other respondents who were not classroom teachers. Almost 96% of the teachers were women, who were distributed across K-3 grade levels. Almost half of the sample reported that they had advanced degrees and had taken an average of 6.6 reading/language arts courses. Nearly one-quarter of the teachers had attended a reading/language arts course within the last year. Teachers reported a wide range of teaching experience in their current grade level ($\underline{M}$=8.6 years) and of total experience ($\underline{M}$=14.8 years). Additional characteristics of the sample are reported in Table 1.

The data were derived from the "CIERA Survey of Early Literacy Programs in High Performing Schools," an instrument created by researchers at the Center for Improvement of Early Reading Achievement in April 1998. The assessment section of the CIERA survey included items arranged in four matrices, to maximize the amount of information obtained from teachers. Each matrix listed a variety of methods to assess children's reading along the left-hand margin, and it required teachers to make judgments about each one according to criteria specified in questions across the top of the page. The topics of the four matrices were: (1) types of assessments and their frequency of use; (2) purposes of assessment; (3) consequences of different assessments; and (4) perceptions of assessment training. For the first item, teachers were provided with six categories of reading assessments (performance assessments, standardized tests, teacher-designed assessments, commercial tests, assessments of fluency and understanding, and assessments of word attack/word meaning), and they were asked to record the specific assessments used in their classroom within the framework provided. Our intent was to provide some structure to the responses but still allow teachers to report the variety of reading assessments that they used. Three blank lines were provided to list the types of assessments for each of the six categories. For example, after "performance assessments," teachers might fill in "running records" and "journal writing" if those were the ways in which they used these assessments in their classrooms. After designating the types of assessments they used in their classrooms, teachers were asked to indicate the frequency with which they used each type of assessment.

Table 1: Teacher Background Characteristics

VARIABLE	N	PERCENT
Gender		
Female	477	96.2
Male	19	3.8
Grade Teaching		
Kindergarten	91	18.1
First Grade	105	20.8
Second Grade	92	18.3
Third Grade	103	20.4
Education Level		
BA	114	22.6
BA+15	139	27.6
MA	128	25.4
MA+15	98	19.4
Ed Specialist	8	1.6
Doctorate	2	.4
Last Course Taken		
Within the last year	121	24.0
1-3 Years ago	146	29.0
3-7 years ago	114	22.6
7+ years ago	104	20.6

What kinds of early reading assessments do teachers use?

High percentages of teachers reported that they used each of the various assessment types; 86% used performance assessments, 82% used teacher-designed assessments, 78% used word attack/word meaning, 74% used measures of fluency and understanding, 67% used commercial assessments, and 59% used standardized reading tests (see Table 2).

Table 2: Examples of Different Types of Assessments

Internal Assessments		External Assessments	
Type	Examples	Type	Examples
Performance assessments	Portfolios Work samples	Standardized reading tests	ITBS Woodcock-Johnson DRP
Teacher-designed assessments	Daily worksheets Observations Anecdotal records	Commercial assessments	Curriculum-based unit tests Workbooks
Fluency and understanding	Assessments of oral reading and comprehension	Word attack and word meaning	Phonemic awareness Vocabulary Sight word assessments

Among teachers who reported using performance assessments, more than 60% reported that they used observations and writing assessments. Of the performance assessments, 22% were observations, 19% were writing assessments, 15% were tests, and 9% were portfolios. For teacher-designed assessments, 34% of the assessments used were observations, 23% informal reading inventories, 19% anecdotal records, 13% work samples, and 10% teacher-designed tests. Note that there is some overlap among categories (e.g., observations and tests). This is due to the open-ended nature of the items, which allowed teachers to determine the category in which to place each of the assessments. Consequently, "observations" were sometimes reported as performance assessments and other times as measures of fluency/understanding. Regardless of such differences in classification, however, it is clear that observations were used very frequently.

For the category of word attack and word meaning, seven different types were reported, including phonics (29%), vocabulary (22%), sight words (19%), tests (12%), oral reading (9%), spelling (5%), and work samples (5%). Teachers reported using the following five types of fluency and understanding assessments: oral reading (43%), comprehension (25%), observations (21%), tests (8%) and work samples (4%).

Commercial and standardized tests—the external assessments least controlled by teachers—were used least often. Teachers indicated that on average they used all categories of assessments, except standardized tests, approximately once per week. They used standardized tests less frequently, only about 2-3 times/year. Teachers reported using the following five types of standardized assessments of reading; norm-referenced (69%), state level (10%), skills (4%), district (2%), curriculum (2%). The remaining 13% of standardized assessments that teachers reported using were idiosyncratic and could not be placed in any specific category. Most of the commercial assessments were workbooks (43%) and basal readers (34%). The remaining types of commercial assessments included curriculum-based (6%), specific program (6%), specific skills (5%), and other types of commercial assessments (6%). In addition to these categories, teachers completed an "other" category where they could list assessments that did not fit within the frame-

work. Although the "other" responses comprised less than 5% of the total responses, most of these were external assessments.

The survey showed that K-3 teachers use a tremendous variety of assessments in their classrooms on a daily basis. Assessments designed by teachers were the most frequently-used type and standardized tests were used least often. This contrast was most evident for K-1 teachers. It may be reasonable to speculate that the trend changes in higher grades, where students usually have more standardized and commercially produced tests. A main finding that emerges from the survey is that K-3 teachers use observations, anecdotal evidence, informal inventories, and work samples as their main sources of evidence about children's reading achievement and progress. A second main finding is the huge variety of tools available to teachers and the large variation in what they use. A high degree of skill is required for a teacher to select and use appropriate assessment tools.

What are the purposes of assessment?

Another matrix in the survey provided teachers with seven different purposes for assessment and asked them to indicate whether they used assessments in each of the six categories for the following purposes: placement, referral, diagnosis, report cards, conferences, summary, and future tests. Because teachers were asked to provide dichotomous responses, in which checkmarks represented the use of the assessment type for each specific purpose, percentages were tabulated which represented the percentage of teachers affirmatively stating that they used a particular assessment category for a specific purpose.

The results showed that teachers used assessments for a variety of purposes, and that some assessment types were used for more purposes than others. Teachers very often used internal assessments of performance, fluency and understanding, and word attack/word meaning for diagnosis, for filling out report cards, and for discussion at parent-teacher conferences. Conversely, few teachers reported using commercial assessments and standardized reading tests for these purposes. Fewer than half of the teachers said that they used commercial assessments for conferences, report cards, and diagnosis. Even fewer teachers said that they used commercial assessments and standardized tests for referrals, conferences, report cards, or placements. Thus, teachers reported using internal assessments more often and for more purposes than they used external assessments.

What is the impact of assessments on various stakeholders?

The next item in the survey, regarding the consequences of assessments, posed seven specific questions about the effects of assessments on various stakeholders. The questions were:

1. How does the assessment affect students' learning to read?
2. How does the assessment affect students' motivation to read?

3. How does the assessment affect the teacher's decisions about what skills and information to teach?

4. How does the assessment affect parents' knowledge about their child's reading performance?

5. How does the assessment affect parents' active involvement in helping their child learn to read?

6. How does the assessment affect administrators' knowledge about students' reading/language arts performance?

7. How does the assessment affect administrators' use of results in public reports?

Teachers answered these questions for each of the six assessment categories (performance, standardized, teacher-designed, commercial, fluency and understanding, word attack/word meaning) using a five-point scale ranging from "1" (strong negative impact) to "5" (strong positive impact). The actual data are shown in Table 3.

Table 3: Mean Impact of Assessment Types on Stakeholders

Types of Assessment	Student Learning	Student Motivation	Teacher Decisions	Parent Knowledge	Parent Involvement	Administrator Knowledge	Administrator Use of Results
Performance assessment	3.8	3.7	4.5	4.0	3.5	4.0	3.7
Standardized reading tests	2.8	2.6	3.7	3.4	3.0	4.1	4.1
Teacher-designed assessments	4.1	3.8	4.7	4.1	3.7	3.6	3.3
Commercial assessments	3.1	3.0	3.7	3.5	3.2	3.6	3.5
Fluency and understanding	4.1	4.0	4.5	3.9	3.6	3.7	3.4
Word attack and word meaning	4.1	3.9	4.5	3.9	3.6	3.7	3.4

With the exception of standardized tests, teachers reported that each of the assessment types had very positive effects on teachers' daily practices in classrooms. Standardized assessments exerted their greatest impact on administrators' knowledge and their use of test results. Teachers reported that assessments designed by teachers, assessments of fluency/understanding, and assessments of word attack/word meaning had the least positive impact on administrators' use of results. Standardized and commercial assessments had the least positive impact on student motivation, and performance assessments had the least positive impact on parent involvement. In general, teachers reported that internal, as compared to external, assessments had more positive effects on students, teachers, and parents. Conversely, teachers believed external assessments had a higher positive impact on administrators. These patterns suggest that teachers differentiate between assessments over which they have control and assessments generated externally, in terms of their impact on stakeholders. It is ironic that teachers believed that the most useful assessments for students, teachers, and parents were valued less by administrators than external assessments.

The study suggests that high-stakes tests do not necessarily mean high benefits for classroom practices and student learning.

Do teachers feel prepared to assess students?

Teachers were asked to indicate how well they believed that they were trained to use each of the six assessment types. They made these judgments based on a scale ranging from "1" (No training) to "5" (Excellent training). On average, teachers reported positive perceptions of training for each of the various assessments. The means, in order of least to most training, were: commercial assessments (M=3.3), standardized tests (M=3.4), performance assessments (M=3.8), fluency and understanding (M=3.9), word attack/word meaning (M=3.9), and teacher-designed assessments (M=4.0). Across all assessment types, teachers reported an overall mean of 3.7 for their perceived level of training on these reading assessments. They rated their training lower or "Fair" on external assessments, and "Good" on the internal assessments. Teachers' perceptions of training adequacy varied as a function of their backgrounds. Teachers with a bachelors degree plus 15 credits reported significantly less training than teachers who either had a masters degree plus 15 credits, a doctoral degree, or an educational specialist degree. Also, teachers who had taken more reading/language arts courses reported that they had better training. The study shows that teachers feel most prepared to use the assessment tools that they create or select and less prepared to use external assessments that are given to them.

Points to Ponder About the Survey

Teachers in the effective schools participating in this study reported using a variety of assessments daily to assess reading. They used many specific reading tests, commercial products, and teacher-designed activities. Indeed, their responses included hundreds of assessments that we grouped according to six types. The types and frequency of assessments varied most for kindergarten teachers, as might be expected, but in general, all teachers in grades 1-3 reported using many kinds of assessments on a weekly basis. Observations and writing were the most frequently mentioned informal and teacher-controlled types of assessments, perhaps because they can be done quickly as part of many curricular activities. Other surveys of teachers' assessment practices and the commercial marketplace of K-3 reading assessments have confirmed the huge variety of tools available to teachers (Meisels, Paris, & Pearson, 1999). Teachers face the formidable task of finding these tools, learning about them, ordering/obtaining them, and then adapting the tools to their own purposes and students.

We noted a contrast between teachers' views of internal and external assessments. Standardized tests and commercial tests that allow little teacher control and adaptation were regarded as less useful and were used less often by

teachers. Teachers also regarded external tests as the least beneficial for students, parents, and teachers. Paradoxically, the external tests were regarded as having the most impact on administrators' knowledge and reporting practices. We think that teachers' frustration with assessments is partly tied to this paradox. Few teachers reported that they had excellent training on any type of assessment; but they rated their training as "Good" for performance assessments and similar teacher-designed assessments, whereas they rated their training on commercial and standardized tests lower. It seems clear that when districts place a premium on the results of external assessments, they need to provide more information and training for teachers on the appropriate use and interpretation of those assessments.

There were few differences among teachers according to teaching experience and educational background. The most frequent effect was for kindergarten teachers, who used assessments less frequently than teachers in higher grades, and who also had more positive perceptions of the impact of assessment on parents' involvement and administrators' knowledge and use. Perhaps kindergarten teachers use assessments primarily for screening, placement, and designing developmentally-appropriate activities, and less for comparative or accountability purposes. The similarity among other teachers in grades 1-3 suggests that they use a variety of internal assessments for similar purposes.

It may not be surprising that successful teachers use assessments that they can design and control more often than "off-the-shelf" tests. Such teachers feel better trained to use these assessments and believe that they have positive benefits for students' learning and motivation, as well as for parental information and involvement. One ironic finding is that the most frequent and beneficial evidence of children's reading may be the least visible and enduring in public reports. Observations, anecdotes, and daily work samples are certainly low-stakes evidence of achievement for accountability purposes, but they may be the most useful for teachers and students. A second irony is that the assessments on which teachers feel least trained and regard as least useful are used most often for evaluations and public reports. Together, these findings suggest that teachers need support in establishing the value of "internal" assessments in their classrooms for administrators and parents, while also demarcating the limits and interpretations of external tests. The current slogan about the benefits of a "balanced" approach to reading instruction might also be applied to a "balanced" approach to reading assessment. The skills that are assessed need to be balanced among various components of reading and the purposes/benefits of assessment need to be balanced among the stakeholders.

The critical question that many policymakers ask is, "Which reading assessments provide the best evidence about children's accomplishments and progress?" The answer may not be one test or even one type of assessment. We know that a single test or assessment cannot represent the complexity of reading. Likewise, one type of assessment may not represent the curriculum and instructional diversity among teachers, nor will the same assessments capture the different skills and developmental levels of children. That is why teachers use multiple assessments, choosing those that fit their purposes and reveal the most information about their students. We believe that the most robust evidence about children's reading reveals developing skills that can be compared to individual standards of progress, as well as normative

standards of achievement. A developmental approach balances the types of assessments across a range of reading factors and allows all stakeholders to understand the strengths and weaknesses of the child's reading profile. Many teachers use this approach implicitly, and we think it is a useful model for early reading assessment.

A Developmental Approach to Assessment

Not many parents or teachers expect assessments to be given to kindergarten children, but such assessments can be very useful. Five- and six-year-olds have emerging knowledge about literacy that varies widely among children depending on their home background and experiences. Early assessments can identify children who know the alphabet, who can write their own name, and who have participated in joint storybook reading—all indicators of rich literacy environments during early childhood. Kindergarten teachers may assess these skills through observation or with brief structured tasks. For example, sharing a book with a child can be an occasion to assess a child's recognition of letters, understanding of print concepts, and ability to retell a sentence or part of the story. For children who cannot identify letters and words, teachers may choose to use wordless picture books to assess knowledge about narratives in connected pictures, a pre-reading skill and a good index of comprehension (Paris & Paris, 2000). Young children's emerging knowledge about letter-sound relations is revealed in their "invented spelling" and can be assessed by teachers who ask children to listen to a dictated sentence and then write it. Each phoneme that a child hears and represents with a letter is an indication that the child is decoding sounds that correspond to distinct letters. Kindergarten teachers can also listen to children "read" familiar books that have been memorized to assess comprehension, accuracy, and word recognition. This is a natural precursor to assessing how children read unfamiliar words and books.

Some children may begin oral reading in kindergarten, but most begin in first grade. Teachers use informal reading inventories (IRIs) to assess oral reading accuracy with running records or miscue analyses. There are commercial IRIs that provide graded word lists, graded passages or leveled books, and directions for administering and scoring them. Whether teacher-designed or commercial, the IRI is a useful task for assessing children's oral reading rate, accuracy, fluency, comprehension, and retelling in a 10–15 minute session. First and second grade teachers weave reading and writing together for both instruction and assessment. For example, they might use a Writers' Workshop activity for children to draw and write about a recent event. They may use process writing in small groups as a means of assessing children's revising skills, while simultaneously encouraging children to read and edit each other's work. Reports, projects, and journals are used frequently in grades 1–3 because children are motivated to write about their own experiences. These work samples, whether assembled in folders, portfolios, or journals, provide excellent assessments of literacy accomplishments that can be shared with children and parents (Paris & Ayres, 1994). Many teachers like to assess children's attitudes about reading and how often they read on their own, so they may ask children to fill out brief sur-

veys, answer open-ended questions, or keep records of when and what they read. Research has shown that young children often read less than 10 minutes per day outside school, and we know that positive attitudes and literacy habits are the foundation for early reading success (Snow, Burns, & Griffin, 1998). Some of the most frequent K-3 literacy assessments are shown below.

Table 4: Common K-3 Assessments

ASSESSMENT TASKS	USUALLY ASSESSED AT GRADE LEVELS			
	K	1	2	3
Letter identification; letter-sound correspondence	X	X		
Phonological awareness (e.g., rhyming, blending)	X	X		
Concepts about print	X	X		
Oral language and listening	X	X		
Decoding and word identification	X	X	X	X
Oral reading rate and fluency		X	X	X
Journals, portfolios, work samples		X	X	X
Comprehension and retelling		X	X	X
Attitudes about reading		X	X	X
Book logs, reading habits, interests			X	X

The battery of assessments shown here is similar to the K-3 assessments included in the Michigan Literacy Progress Profile (MLPP, 2000) designed by the state Department of Education and Michigan educators. The MLPP is intended to be a resource which teachers use selectively—with some of their students some of the time, rather than with all students. The state legislature has recommended that the MLPP can be used to assess annual student progress as well as achievement in summer school programs. Thus, it is a hybrid assessment that has features of both internal assessments (e.g., teacher control) and external assessments (e.g., uniformity and external credibility). Other states are developing similar early assessment tools. For example, Texas has created the Texas Primary Reading Inventory (TPRI) for teachers to use as an assessment tool with K-3 children.

One key to a developmental approach to assessment is matching the battery of assessments to the child's emerging abilities, so that teachers and parents understand the child's strengths and weaknesses. Teachers need to be aware of the many assessment tasks in order to choose them appropriately, and the number and variety of assessments is daunting. A second key to assessment is keeping records of progress with multiple assessments throughout the year, so that each child's development can be recorded and interpreted. A third point is that assessment should occur daily and be integrated with instruction, in order for teachers to provide instruction that is challenging and appropriate for each child. These and other practices were noted by the National Association for the Education of Young Children (NAEYC) in a position statement in 1990. Their guidelines are summarized below.

Guidelines for Assessment with Children 3–8 Years of Age (NAEYC, 1990)

1. Curriculum and assessment are integrated throughout the program.
2. Assessment results in benefits to the child.
3. Children's development in all domains is assessed informally and routinely by teachers' observations.
4. Assessment provides teachers with useful information to successfully fulfill their responsibilities.
5. Assessment involves regular and periodic observations of children in a wide variety of circumstances.
6. Assessment relies on procedures that reflect typical activities in the classroom.
7. Assessment relies on demonstrated performance during real, not contrived, activities.
8. Assessment utilizes an array of tools and a variety of processes.
9. Assessment recognizes individual diversity of learners and allows for differences in styles and rates of learning.
10. Assessment supports children's development; it does not threaten their safety or self-esteem.
11. Assessment supports parents' relationships with their children; it does not undermine parents' confidence in their children nor devalue the language and culture of the family.
12. Assessment demonstrates children's overall strengths and progress.
13. Assessment is an essential component of the teacher's role.
14. Assessment is a collaborative process involving children and teachers, teachers and parents, and school and community.
15. Assessment encourages children to participate in self-evaluation.
16. Assessment demonstrates what children can do independently and what they can do with assistance.
17. Information about each child's development is systematically collected and recorded at regular intervals.
18. A regular process exists for periodic information sharing between teachers and parents.

So Many Children, So Little Time

Using effective reading assessment is not easy. Teachers often complain that it takes too much time to assess children individually on a regular basis. They also say that the wide range of reading abilities in their classrooms makes assessment difficult. Even when they can administer reading assessments, teachers report that it is difficult for them to interpret the results in a straightforward way for children and parents. Paradoxically, the internal assessments are regarded as having low stakes by administrators. Consequently, teachers may feel frustrated that no one cares how well they use informal assessments. Anyone who works in schools knows these problems firsthand. There is no simple solution, but we have seen how effective teachers deal with these issues. Here are some tips from effective assessment practices that we have observed skilled teachers using in many schools:

1 Effective assessment must involve the whole school or district. The system gains visibility and credibility among parents when it is endorsed by the entire school. Teachers cannot create assessment systems alone; they need emotional support, collaborative teamwork , and shared motivation to build a system of assessment tools that serves their students and community. Effective schools have principals and administrators who provide leadership in integrating assessment with the curriculum and instruction.

2. Literacy assessments cannot be built piecemeal. Effective reading assessment is a woven fabric, not a tangle of individual threads. It starts with a shared conception of the goals of reading assessment and the variety of tools that are available. Then it proceeds to the sharing of effective practices. Teachers need to choose a coherent set of assessment tools that can be used on a regular basis. State-level assessment batteries or standards, district outcomes or portfolios, and school-based report cards are all options in a systematic approach to reading assessment. The system itself indicates to students, parents, and the public what is important to master in the curriculum.

3. Sustained professional development activities are supported by outside experts and internal teams of teachers, who design and revise their assessment system. It may take several years for all teachers to understand assessment practices and use them in similar ways. Consensus is built upon regular reflection and discussion among teachers about what assessments are working well, how the assessments support parent conferences and report cards, and how assessments help individual children.

4. Effective reading assessment means using assessments selectively. Too many districts add new assessments on top of old ones, burdening both teachers and students with redundancies. Choices must be made about which assessments are used with which students. Teachers may choose to assess their lowest-achieving readers more often or more thoroughly than children reading at or above grade level. They may choose to use different types of assessment with children who are reading poorly or well, so that every child can demonstrate his or her best accomplishments. Not all children need to be given the same assessments.

5. Teachers use assessments to reinforce their professional evaluations of children for report cards and parent conferences. Because the assessment process is an ongoing part of teaching in the classroom, it should draw on multiple sources of evidence and should produce a variety of mutually-supporting reports. Assessments of students are important public reports that support teachers' professional judgments about students. Parents need to understand the kinds of evidence used to evaluate student achievement and need to trust teachers' interpretations of the evidence. As the public becomes better informed about literacy assessment and the evidence it provides about children, parents may gain more confidence in internal assessments and lose some of their confidence in external assessments. Such views are shared by many teachers.

Conclusions

Assessment is becoming increasingly important for teachers in primary grades because administrators and parents want more detailed information about children's early literacy achievement and progress. Yet teachers believe that their primary mission is instruction and support for the child's whole development. Many teachers, frustrated by the pressure to assess and report results of tests that they feel provide only partial or fragile evidence, resist spending time on assessments, especially if they are for external purposes. The public needs to understand the difficulties and limitations of early assessments and the need for multiple sources of evidence. Many teachers prefer to rely on professional judgment, supported by prudent use of various literacy assessments, feeling that this approach is more beneficial for children and parents.

Effective assessment does not mean simply training teachers to use new tests wisely, although such expertise is important. Assessment reform in schools must also involve communication and negotiation among stakeholders about the kinds of information that support students' educational growth. The CIERA survey revealed that teachers perceive large differences between administrators' value and use of assessment information and those of other stakeholders. Administrators (and parents) need to learn how teachers use reading assessments, just as much as teachers need to learn new kinds of assessments.

The CIERA survey confirms and extends our understanding of effective assessment practices with young children. At the simplest level, assessment should be a way to communicate information about children's accomplishments. If children's welfare is the highest educational priority, then teachers, parents, and administrators should work together to design assessment systems that bring the greatest benefits to children. We believe that a developmental approach to assessment is part of this solution. It is not a one-size-fits-all approach, nor an approach that gives the same test to all children on the same day. Instead, assessment is embedded in daily classroom activities, in which teachers use formal and informal assessment tools to ascertain if children are improving their literacy skills and knowledge, mastering the curriculum, and meeting community standards of literacy development. These practices are effective because they empower teachers and students alike.

References

American Educational Research Association (2000). Position statement of the American Educational Research Association concerning high-stakes testing in preK–12 education. *Educational Researcher, 29*(8), 24–25.

Haertel, E. (1989). Student achievement tests as tools of educational policy: Practices and consequences. In B.R. Gifford (Ed.) *Test policy and test performance: Education, language, and culture* (pp. 25–50). Boston: Kluwer Academic Publishers.

Haladyna, T., Nolen, S. B., & Haas, N. S. (1991). Raising standardized achievement test scores and the origins of test score pollution. *Educational Researcher, 20*, 2–7.

Hoffman, J. V., Assaf, L., Pennington, J. & Paris, S. G. (2001). High stakes testing in reading: Today in Texas, tomorrow? *The Reading Teacher, 54*(5), 482–492.

International Reading Association (1999). *High-stakes testing in reading. A position paper.* Newark, DE: International Reading Association.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4–16.

Madaus, G.F., & Tan, A.G.A. (1993). The growth of assessment. In G. Cawelti (Ed.), *Challenges and achievements of American education* (pp. 53–79). Alexandria, VA: Association for Supervision and Curriculum Development.

Meisels, S., Paris, S. G., & Pearson, P. D. (1999). *Three perspectives on early reading assessment: What do we use? What do we sell? What are their consequences?* Symposium presented at the National Conference on Large-Scale Assessment, Snowbird, Utah.

Michigan Literacy Progress Profile. (2000). Lansing, MI: Department of Education.

Murphy, S., Shannon, P., Johnston, P., & Hansen, J. (1998). *Fragile evidence: A critique of reading assessment*. Mahwah, NJ: Lawrence Erlbaum Associates.

National Association for the Education of Young Children (1990). Guidelines for appropriate curriculum content and assessment in programs serving children ages 3 through 8. Washington, DC: NAEYC.

National Commission on Testing and Public Policy (1990). *From gatekeeper to gateway: Transforming testing in America*. Chestnut Hill, MA: National Commission on Testing and Public Policy.

Nolen, S. B., Haladyna, T. M., & Haas, N. S. (1992). Uses and abuses of achievement test scores. *Educational Measurement: Issues and Practices, 11*, 9–15.

Paris, S. G. (2000). Trojan horse in the schoolyard: The hidden threats in high-stakes testing. *Issues in Education, 6*(1,2), 1–16.

Paris, S. G., & Ayres, L. J. (1994). *Becoming reflective students and teachers with portfolios and authentic assessment*. Washington, DC: American Psychological Association.

Paris, A. H., & Paris, S. G. *Innovative comprehension assessments for non-readers*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, April 2000.

Paris, S. G., Roth, J. L., & Turner, J. C. (2000). Developing disillusionment: Students' perceptions of academic achievement tests. *Issues in Education, 6*(1,2), 17-45.

Paris, S. G., Turner, J. C., Lawton, T. A. & Roth, J. L. (1991). A developmental perspective on standardized achievement testing. *Educational Researcher, 20*, 12-20.

Paris, S. G., & Urdan, T. (2000). Policies and practices of high-stakes testing that influence teachers and schools. *Issues in Education, 6*(1,2), 83-107.

Roth, J., Paris, S. G., & Turner, J. C. (2000). Students' perceived utility and reported use of test-taking strategies. *Issues in Education, 6*(1,2), 67-82.

Shepard, L. (1991). Interview on assessment issues with Lorrie Shepard. *Educational Researcher, 20*, 21-23.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher, 20*, 8-11.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in children*. Washington, DC: National Academy Press.

Urdan, T. C., & Paris, S. G. (1994). Teachers' perceptions of standardized achievement tests. *Educational Policy, 8*, 137-156.

Wong, C.A., & Paris, S.G. (2000). Students' beliefs about different kinds of tests. *Issues in Education, 6*(1,2), 47-66.

About CIERA

The Center for the Improvement of Early Reading Achievement (CIERA) is the national center for research on early reading and represents a consortium of educators in five universities (University of Michigan, University of Virginia, and Michigan State University with University of Southern California and University of Minnesota), teacher educators, teachers, publishers of texts, tests, and technology, professional organizations, and schools and school districts across the United States. CIERA is supported under the Educational Research and Development Centers Program, PR/Award Number R305R70004, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

Mission. CIERA's mission is to improve the reading achievement of America's children by generating and disseminating theoretical, empirical, and practical solutions to persistent problems in the learning and teaching of beginning reading.

CIERA Research Model

The model that underlies CIERA's efforts acknowledges many influences on children's reading acquisition. The multiple influences on children's early reading acquisition can be represented in three successive layers, each yielding an area of inquiry of the CIERA scope of work. These three areas of inquiry each present a set of persistent problems in the learning and teaching of beginning reading:

CIERA INQUIRY 1
Readers and Texts

Characteristics of readers and texts and their relationship to early reading achievement. What are the characteristics of readers and texts that have the greatest influence on early success in reading? How can children's existing knowledge and classroom environments enhance the factors that make for success?

CIERA INQUIRY 2
Home and School

Home and school effects on early reading achievment. How do the contexts of homes, communities, classrooms, and schools support high levels of reading achievement among primary-level children? How can these contexts be enhanced to ensure high levels of reading achievement for all children?

CIERA INQUIRY 3
Policy and Profession

Policy and professional effects on early reading achievement. How can new teachers be initiated into the profession and experienced teachers be provided with the knowledge and dispositions to teach young children to read well? How do policies at all levels support or detract from providing all children with access to high levels of reading instruction?

www.ciera.org

CIERA
Center for the Improvement of Early Reading Achievement

CIERA is a collaboration of
University of Michigan
University of Virginia
Michigan State University
with
University of Minnesota
University of Southern California

University of Michigan School of Education
610 E University Av, Rm 1600 SEB
Ann Arbor, MI 48109-1259
734.647.6940 voice
734.763.1229 fax
ciera@umich.edu

www.ciera.org

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

REPRODUCTION BASIS

☐ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☑ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").